Finding “Interesting” Correlations in Multi-Faceted Personal Informatics Systems

Simon L Jones
University of Bath
Bath, BA2 7AY, UK
s.l.jones@bath.ac.uk

Ryan Kelly
University of Bath
Bath, BA2 7AY, UK
r.m.kelly@bath.ac.uk

CHI'16 Extended Abstracts, May 07-12, 2016, San Jose, CA, USA
ACM 978-1-4503-4082-3/16/05.
http://dx.doi.org/10.1145/2851581.2892401

Abstract

Personal informatics systems are capable of uncovering interesting insights about their users by identifying statistical correlations in multi-faceted data. However, they often produce an overwhelming quantity of information. We explore the feasibility of automatically filtering correlational information based on its interest to users. We analyze users’ subjective ratings of correlations in their data to gain a deeper understanding of the factors that contribute to users’ interest. We then use this understanding to identify candidate objective measures for information filtering, which can be applied without input from the user. Finally, we test the predictive power of these measures. Our main findings reveal that users in our study valued the *Surprisingness*, *Utility* and *Positive Valence* of correlational information above other factors.

Author Keywords

Personal Informatics; Data mining; Information Filtering; Quantified Self; Information Dashboards.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

Data Category	Attribute
Day	day_of_week
Events	events events_duration
Location	checkins location
Mood	mood_score mood_note
Music	tracks
Physical Activity	steps steps_active_min steps_distance
Productivity	distracting_min neutral_min productive_min
Sleep	sleep sleep_awakenings sleep_end sleep_start time_in_bed
Social Media	instagram_comments instagram_likes instagram_posts tweets twitter_mentions
Weather	weather_cloud_cover weather_precipitation weather_temp_max weather_temp_min weather_wind_speed

Table 1: Data Categories and Attributes tracked by Exist.io, all of which are included in the pair-wise correlation analysis.

Introduction

Personal informatics systems have been shown to provide value in a variety of settings, from supporting reminiscence [4][5] to managing chronic medical conditions [9][11]. Many systems have the capability to uncover associations between distinct aspects of a person's life (e.g. physical activity and sleep) through analysis of personal tracking data [1]. A growing number of mainstream personal informatics tools are adopting this multifaceted approach to personal tracking. Examples include Exist[1], TicTrac[2], and Zenobase[3]. These systems apply statistical analysis to explore the relationships among diverse forms of personal data. The objectives of such systems are to simplify the management and reduce the workload of analyzing data by processing it within a single repository [2]. An additional aim is to provide users with holistic insights that they could not easily derive themselves.

However, the output produced by automated analysis tools may be unmanageable for users [10]. For example, a tool that explores pairwise correlations among 20 multifaceted variables has the potential to report up to 190 relationships [10]. We argue that there is likely to be a significant risk of information overload associated with making sense of so many novel and potentially unanticipated observations. There is thus a need to understand how we can support users in making sense of this data and how we can help them to identify insights that are of real interest.

A significant challenge, however, is that it is not clear what outputs users deem to be 'interesting'. Previous studies have shown that some results of automated analysis are considered more interesting than others, and that 'obvious' observations are of little value to users [1][10]. Measuring the '*Interestingness*' of patterns within data is an important area of data mining research [7], and long-term engagements with personal tracking systems are predicated on the extent to which they can inform users about interesting patterns [12]. But automatically evaluating and filtering results in personal informatics systems, in order to present those that are of most interest to the user, remains an open challenge.

In this paper we investigate users' interest in the correlational information presented by a personal informatics system and the characteristics of data that might predict this interest, as a first step towards automated filtering. We present late-breaking results from a study in which 18 participants used Exist.io[1], a commercial personal data aggregation and analysis system, for periods lasting between 1-4 months. In the first part of this paper we identify subjective measures associated with *Interestingness* of correlations. In the second part of the paper this informs the selection of objective measures that may support information filtering in personal informatics systems.

Exist.io User Study

Exist allows users to combine data from numerous distinct self-tracking services and explore statistical correlations present within their data. The Exist platform advertises itself as a tool to "understand your life[1]" and, like many other personal informatics systems, enables exploratory use, whereby users can volunteer as much data as possible in order to see if interesting insights emerge.

[1] http://exist.io
[2] http://www.tictrac.com
[3] http://www.zenobase.com

Participants were recruited via an advertisement placed on an online University noticeboard. Applicants were screened and selected on the basis of having a general interest in the use of self-tracking technologies.

Participants were entered into a prize draw for a single £50 Amazon.co.uk voucher as an additional incentive to participate in the study.

At the time of writing, 13 (of 18) participants had completed the study, providing daily tracking data to Exist.io for 1-4 months. These participants consisted of 6 males and 7 females, with an age range of 25-59yrs (Mean age = 33yrs).

3 of the participants had previous experience of using wearable fitness tracking devices. 2 other participants already tracked their music listening habits using Last.fm.

At the end of the study participants were presented with the correlational information output from Exist.io and interviewed about their reactions to the correlations. Participants were also given a random sample of ten statistically significant correlations within their data, which they rated according to the subjective measures shown in Table 3.

Table 2: Study design and participant information.

At the point of recruitment, all of the participants in our study expressed curiosity and interest in using tracking technologies to uncover information about themselves, but did not have clearly defined problems to address, nor hard-set goals for behaviour change. We argue that these participants reflect a growing proportion of personal informatics users, following the emergence of personal tracking as an everyday activity [8][14].

For example, participants in our study described their motivations for participating in the following ways: “I’m interested to see what it [Exist] tells me about myself” (Participant 4), “I want to know what tracking my life can offer me” (Participant 5), and “Exist seems like a *cool* system. I just want to give it a try” (Participant 12). The results presented in this paper are based on data from 13 participants. The remaining participants had not completed the study at the time of writing. Further details about the participants and the study design are given in Table 2.

Participants provided Exist with data that included daily measures of: physical activity and sleep (both recorded by a wearable Fitbit sensor); productivity and distracting time (recorded by RescueTime logging software); mood (self-reported Likert-scale scores by daily emails); events (automatically retrieved from online calendars); social media interactions (from Twitter and Instagram); music listening (recorded by Last.fm 'scrobbling' from music players such as Spotify and iTunes); and local weather conditions (from Forecast.io). Table 1 shows the full list of data categories and attributes for the data collected from participants.

For each pairwise combination of attributes, the data provided to Exist is analysed for linear correlations. The service then presents correlational information to its users, in the form of graphical visualisations and natural language statements, e.g. ‘You sleep better on days when you are more physically active’, or ‘You are more productive when you listen to classical music’.

In the following sections we explore the subjective factors that affect users’ interest in the correlations presented to them by the Exist system, with the goal of informing our approach to automated filtering of correlational information using objective data features.

What Makes Correlational Information Interesting?

Interestingness measures play an important role in data mining when there is a need to select or rank patterns according to their relevance to the user and reduce information overload [7]. In the context of multi-faceted personal informatics systems, however, it remains unclear what constitutes an ‘interesting’ correlation, particularly when users lack precise goals.

Geng and Hamilton [7] argue that *Interestingness* is best treated as a broad concept that encapsulates a variety of measures. To identify the measures associated with *Interestingness* of correlational information in Exist, we asked participants to rate a random sample of ten statistically significant correlations within their data, according to ten statements (see Table 3). These statements were generated based on existing measures of *Interestingness* from data mining research (e.g. Utility, Surprisingness, Novelty, Actionability) [7], and on our own discussions about the factors that might influence users’ assessments of the correlations, namely; Valence (i.e. the intrinsic attractiveness (positive valence) or aversiveness (negative valence) of a correlation), Uniqueness, and Accuracy/Reliability.

Interestingness (Dependent Variable) - This is a correlation that I would be interested to see when using this system…
Accuracy/Reliability – This is a correlation that is accurate/reliable…
Actionability - This is a correlation that I could use to take action…
Novelty - This is a correlation that tells me something new about myself…
Stability - This is a correlation that is likely to change in different situations/contexts…
Surprisingness - This is a correlation that is surprising/ unexpected…
Uniqueness - This is a correlation that makes me unique compared to other people…
Utility - This is a correlation that is useful to me…
Valence (Positivity) - This is a correlation that is positive/ pleasing …
Valence (Negativity) - This is a correlation that is negative/ displeasing…

Table 3: Participants were asked to respond to each of the above statements on a Likert Scale ranging from 1–Strongly Disagree to 5–Strongly Agree

Responses to these statements, on a Likert Scale ranging from 1–Strongly Disagree to 5–Strongly Agree, captured users' perceptions (subjective measures) of the correlations, rather than actual characteristics of the data (objective measures). The following regression analysis explores which of these subjective measures are associated with *Interestingness*.

Regression Analysis: Predicting Interestingness from Subjective Measures

We conducted a multiple linear regression analysis (using the 'Enter' method) to examine the relationship between overall *Interestingness* (DV), and *Utility*, *Surprisingness*, *Positivity*, *Negativity*, *Uniqueness*, *Novelty*, *Stability*, *Actionability* and *Reliability* (IVs) of correlations. The regression model revealed that *Utility* (β=0.370, $p<.01$), *Surprisingness* (β=0.378, $p<.01$) and *Positive Valence* (β=0.221, $p<.05$) were significant predictors of *Interestingness*. *Actionability* (p=.211), , *Negative Valence* (p=.070), *Uniqueness* (p=.639), *Stability* (p=.429), *Novelty* (p=.843) and *Reliability* (p=.117) were not significant predictors. The overall regression model fit was R^2=0.679, RMSE=0.807.

Regression Model: *Interestingness = .804 + .370*Utility + .378*Surprisingness + .221*PositiveValence*

Discussion of Results

Our analysis reveals that participants were most interested to see correlational insights that were (a) unexpected, e.g. those which presented contradictions of existing knowledge, (b) which offered some utility for understanding or controlling an aspect of their behavior, e.g. learning that their mood was associated with the amount of music that they listened to, and (c) which revealed positively valence, pleasing behaviours. It is worth noting that negatively valenced correlations were approaching statistical significance (p = 0.07), which may suggest that users are interested in uncovering trends that are strongly valenced in either direction (i.e. favouring pleasing, positive trends, and displeasing, negative trends over neutral trends).

Our results support previous findings, which suggest that personal informatics systems should be cautious about simply filtering information according to stated goals [6]. Our participants were frequently interested by results they had not anticipated because they learned something surprising and unexpected. One implication of users' positive reactions to results they were not expecting to see might be that there is value in presenting insights that are unrelated to stated goals or expected areas of importance.

We found that the accuracy of correlations was not significantly associated with overall *Interestingness*. While participants were sometimes cautious about trusting the results produced by Exist (e.g. due to perceived inaccuracies with the tracking technologies, or sparse data), accuracy appeared to be a secondary concern. Participant 1 said: "this correlation is really interesting, *but* I don't know how much I trust it". Interesting results that raised questions regarding accuracy were often a stimulus for further investigation. For example, when viewing a potentially spurious correlation between weather and productivity, one participant told us: "That's fascinating…I'd like to collect more data to see if the correlation stays the same" (Participant 9). Our results suggest that users found value in correlations that generated interesting hypotheses for further investigation, rather than just those that gave definitive results.

In the following sections we turn our attention to objective measures of the data, which might reflect the *Surprisingness*, *Utility and Positive Valence* of

For the following measures an *outcome* is defined as being either a Positive correlation, Negative correlation, or No correlation between a pair of attributes, A and B.

***1. Generality/Coverage*$_{A,B}$ =**

$$\frac{Max(Number\ of\ users\ with\ outcome_{A,B})}{Number\ of\ users\ with\ any\ outcome_{A,B}}$$

Correlations between A and B are general if many users have the same outcome.

***2. Diversity*$_{A,B}$ =**

$$\frac{\left(\begin{array}{c}Number\ of\ distinct\ outcomes\\ for\ all\ users\end{array}\right)_{A,B}}{Number\ of\ possible\ outcomes\ (3)_{A,B}}$$

Correlations between A and B are diverse if we observe all possible outcomes within the data (i.e. positive, negative, no correlation).

***3. Peculiarity*$_{A,B}$ =**

$$1-\frac{\left(\begin{array}{c}Number\ of\ users\ with\ same\\ outcome\ as\ \ selected\ user\end{array}\right)_{A,B}}{Number\ of\ users\ with\ any\ outcome_{A,B}}$$

A correlation is peculiar if few users receive the same outcome as a particular user.

4. *Data Category*: The category to which the data belongs: Day/Events/Mood/Music/Physical Activty/Productivity/Sleep/Social Media/Weather.

5. *Uni/Multi-faceted Correlation*: Correlations between attributes within the same data category are uni-faceted. Correlations between attributes from different data categories are multi-faceted

Table 4: Objective measures of correlation data.

correlations and support automated filtering of correlational information.

Automated Filtering of Correlations

One common approach to address the problem of information overload is to automatically select information that is most likely to be of interest to users [7]. To do this, it is necessary to identify predictors that are available without requiring significant effort from the user. Hence, predictors ought to be objective, measurable characteristics of data that are already available to the system.

A significant challenge is that *Utility*, *Surprisingness* and *Positive Valence* are subjective measures, which take into account factors associated with the user, as well as the data. Previous research emphasizes that subjective measures require access to the user's background knowledge about the data [7]. Requiring users to represent their existing knowledge is a complex and burdensome task.

As a substitute for finding *surprising* results from comparisons between data and existing knowledge, we hypothesize that it may be useful to compare the correlations found for a particular user with those of other users. For example, correlations that are rarely detected for *any* user may be surprising if they appear for an individual. Although users of Exist lack visibility of other users' correlations, we posit that they may have an instinctive sense of which correlations are more or less likely to be prevalent.

We identify three objective measures based on comparisons of correlations between users, namely: *Generality* (how common a correlation is amongst all users), *Diversity* (the diversity in observed outcomes for a particular correlation throughout the entire user population), and *Peculiarity* (the distance of a given correlation from the rest of the data). More detailed definitions are given in Table 4.

With regards to finding *useful* and *positively valenced* correlations, we hypothesize that correlations associated with certain data categories (e.g. Mood, Physical Activity, Sleep etc.) may be inherently more useful or pleasing than others. For example, users may feel that learning about factors associated with their mood outstrips the utility and gratification of learning about factors associated with their social media use. Hence, we select *Data Category* as an objective measure that may have some potential for predicting *Interestingness*. Furthermore, we hypothesize that multi-faceted correlations (inter-correlations between two distinct data categories) e.g. Physical Activity vs. Sleep, are likely to be more useful than uni-faceted correlations (intra-correlations between attributes from the same category), e.g. Sleep: time to bed vs. Sleep: time spent asleep. Previous studies have shown that users are often interested in insights that span multiple different types of data [6][13], and that multi-faceted analysis encourages engagement with personal informatics technologies [1].

The following regression analysis explores whether any of the selected objective measures (shown in Table 4) are predictive of *Interestingness*.

Regression Analysis: Predicting *Interestingness* from Objective Measures

A multiple linear regression analysis was conducted (using the 'Enter' method) to examine the relationship between *Interestingness* (DV) and objective measures of the correlation data, namely: *Generality/Coverage*, *Diversity*, *Peculiarity*, *Data Category* and *Uni/Multi-faceted Correlation* (IVs). The regression analysis revealed that *Uni/Multi-faceted Correlation* (β=-0.511, p<.01) was the only significant predictor of

Interestingness. The overall regression model fit was R^2=0.307, RMSE=1.181.

Regression Model: *Interestingness = 2.232 - 0.511*Uni/Multi-faceted*

Peculiarity/Distance (p=.715), *Generality/Coverage* (p=.575), *Diversity* (p=.293) and all *Data Categories*: Day (p=.596), Events (p=.186), Mood (p=.676), Music (p=.137), Physical Activity (p=.221), Productivity (p=.275), Sleep (p=.118), SocialMedia (p=.856), Weather (p=.811) were not significant predictors of *Interestingness*.

Discussion of Results

Previous work by Bentley et al. [1] reported that obvious correlations were of little value to users. Our results reveal that correlations involving attributes from the same facet or data category are of less interest than those from different categories. We believe that uni-faceted insights reflect many of the obvious correlations previously highlighted as being of little interest to users, owing to the fact that closely related aspects of the same facet are often co-linear, e.g. steps taken vs. distance travelled. Designers of personal informatics systems should take this into account, possibly by demoting uni-faceted observations that are less likely to be of interest. For many of the participants in our study, the removal of uni-faceted correlations would reduce the number of correlations presented by approximately 32%. Yet the imperfect fit of our regression model suggests that a small proportion of uni-faceted correlations are interesting, meaning that further work is required to find ways to automatically identify these exceptions to the rule.

Our results show that the measures based on comparisons of correlations between individual users and the wider populations of users were not effective for identifying interesting information. Therefore, alternative approaches for reducing the overwhelming amount of information that personal informatics systems generate should be explored. For example, it may be necessary to adopt a semi-automated filtering approach, which can account for some aspects of users' existing knowledge and involve the user in the filtering process [7], or a collaborative filtering approach, which automates the process of sharing opinions on the interestingness or relevance of information and requires relatively little effort from the user [2].

Conclusion

We have shown that participants who were curious to investigate the value that a multi-faceted personal informatics system could provide were most interested by correlational insights that were considered surprising, useful, and/or pleasing. This leads to the implication that filtering mechanisms should employ *Surprisingness*, *Utility* and *Valence* as foremost criteria when attempting to reduce the amount of information shown to users. However, we found that several objective measures of the data, which we posit as being related to *Surprisingness, Utility* and *Valence*, did not offer significant value for automatically predicting *Interestingness* of correlations. It is important to note that our search for objective measures was not exhaustive and thus this is an area for future work. Nevertheless, our analysis does reveal that ranking multi-faceted correlations above uni-faceted correlations is an effective approach for prioritising correlational information that is more likely to be of interest to users. We are currently recruiting additional participants for this study in order to strengthen the generalisability and reliability of our results.

References

1. Frank Bentley, Konrad Tollmar, Peter Stephenson, Laura Levy, Brian Jones, Scott Robertson, Ed Price, Richard Catrambone, and Jeff Wilson. 2013. Health Mashups: Presenting Statistical Patterns between Wellbeing Data and Context in Natural Language to Promote Behavior Change. *ACM Transactions of Computer-Human Interaction*. 20, 5, Article 30 (November 2013), 27 pages. http://dx.doi.org/10.1145/2503823

2. Al Borchers, Jon Herlocker, Joseph Konstan, and John Riedl. 1998. Ganging up on Information Overload. *Computer.* 31, 4, (April 1998), 106-108. http://dx.doi.org/10.1109/2.666847

3. Eun Kyoung Choe, Nicole B. Lee, Bongshin Lee, Wanda Pratt, and Julie A. Kientz. 2014. Understanding quantified-selfers' practices in collecting and exploring personal data. In Proceedings of the 32nd annual ACM conference on Human factors in computing systems (CHI '14). ACM, New York, NY, USA, 1143-1152. http://doi.acm.org/10.1145/2556288.255737

4. Dan Cosley, Kathy Akey, Brian Alson, Jonathan Baxter, Mark Broomfield, Soyoung Lee, and Chethan Sarabu. 2009. Using technologies to support reminiscence. In Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology (BCS-HCI '09). British Computer Society, Swinton, UK, UK, 480-484.

5. Chris Elsden and David S. Kirk. 2014. A quantified past: remembering with personal informatics. In Proceedings of the 2014 companion publication on Designing interactive systems (DIS Companion '14). ACM, New York, NY, USA, 45-48. http://dx.doi.org/10.1145/2598784.260277

6. Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming data complexity in lifelogs: exploring visual cuts of personal informatics data. In Proceedings of the 2014 conference on Designing interactive systems (DIS '14). ACM, New York, NY, USA, 667-676. http://dx.doi.org/10.1145/2598510.2598558

7. Liqiang Geng and Howard J. Hamilton. 2006. Interestingness measures for data mining: A survey. *ACM Computing Surveys*. 38, 3, Article 9 (September 2006). http://dx.doi.org/10.1145/1132960.1132963

8. Cathal Gurrin, Alan F. Smeaton, and Aiden R. Doherty. 2014. LifeLogging: Personal Big Data. *Foundations and Trends in Information Retrieval*. 8, 1 (June 2014), 1-125. http://dx.doi.org/10.1561/1500000033

9. Kit Huckvale, Mate Car, Cecily Morrison and Josip Car. 2012. Apps for asthma self-management: a systematic assessment of content and tools. BMC medicine, 10, 144. DOI=10.1186/1741-7015-10-144

10. Simon L. Jones. 2015. Exploring correlational information in aggregated quantified self data dashboards. In Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct). ACM, New York, NY, USA, 1075-1080. http://dx.doi.org/10.1145/2800835.2800963

11. Ravi Karkar, James Fogarty, Julie A. Kientz, Sean A. Munson, Roger Vilardaga, and Jasmine Zia. 2015. Opportunities and challenges for self-experimentation in self-tracking. In Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers (UbiComp/ISWC'15 Adjunct). ACM, New York, NY, USA, 991-996. http://dx.doi.org/10.1145/2800835.2800949

12. Amanda Lazar, Christian Koehler, Joshua Tanenbaum, and David H. Nguyen. 2015. Why we use and abandon smart devices. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '15). ACM, New York, NY, USA, 635-646. http://dx.doi.org/10.1145/2750858.2804288

13. Ian Li, Jodi Forlizzi, and Anind Dey. 2010. Know thyself: monitoring and reflecting on facets of one's life. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '10). ACM, New York, NY, USA, 4489-4492. http://dx.doi.org/10.1145/1753846.1754181

14. Melanie Swan. 2012. Health 2050: The realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen. *Journal of Personalized Medicine*. 2, 3, 93–11